

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 7, 2005, 07:04:17 ; Search time 8.94985 Seconds

(without alignments)  
1193.323 Million cell updates/sec

Title: US-09-939-537-37

Perfect score: 591

Sequence: 1 TRFSRSABPPAYQGQGNOLY.....LSTATKDTYDALHMQALPPR 111

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

PIR 79: \*  
1: p1r1: \*  
2: p1r2: \*  
3: p1r3: \*  
4: p1r4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	583	98.6	163	2 A31768	T-cell receptor ze
2	476.5	80.6	164	2 A40104	T-cell receptor CD
3	471.5	79.8	166	2 JC4664	T-cell receptor ze
4	471.5	79.8	166	2 A46424	T-cell surface gly
5	389	65.8	188	2 A45089	T-cell receptor CD
6	378.5	64.0	177	2 S54817	T-cell receptor id
7	378.5	64.0	206	2 A35900	T-cell receptor CD
8	86	14.6	1154	2 S69206	regulator protein
9	80	13.5	692	2 T21095	hypothetical prote
10	76.5	12.9	527	2 S38329	probable ATP-depen
11	74	12.5	136	2 T14305	probable late embr
12	74	12.5	3190	2 T13828	CREB-binding prote
13	73.5	12.4	123	1 S26183	general stress pro
14	73.5	12.4	360	2 AF0317	rare lipoprotein A
15	73.5	12.4	374	2 H70415	conserved hypochet
16	73.5	12.4	1258	2 T30252	nuclear protein SA
17	71.5	12.1	632	2 T02627	hypothetical prote
18	70.5	11.9	594	2 G96525	protein TIM5.25 l
19	70	11.8	193	2 AG2567	hypothetical prote
20	70	11.8	448	2 A36311	70K UI small nucle
21	70	11.8	568	1 I61106	evthiocytin - rat
22	70	11.8	2441	2 D71623	hypothetical prote
23	69.5	11.8	185	2 S71512	hypothetical prote
24	69.5	11.8	205	2 T20962	hypothetical prote
25	69.5	11.8	1269	2 F84730	probable myosin he
26	69	11.7	393	2 T05532	hypothetical prote
27	69	11.7	488	2 G81213	conserved hypochet
28	69	11.7	490	2 C81790	conserved hypochet
29	68.5	11.6	615	2 G87678	ATP-dependent DNA

30	68	11.5	123	2 S07364	seminal vesicle se
31	68	11.5	251	2 T17784	PBCV-1 33kd peptid
32	68	11.5	319	2 I50519	sox 19 protein - z
33	68	11.5	384	1 A43710	involucrin - chick
34	67.5	11.4	212	2 S40027	protein p71 - mous
35	67.5	11.4	249	2 P95957	probable transcrip
36	67.5	11.4	761	2 F82078	topoisomerase IV,
37	67.5	11.4	784	2 G95112	exoribonuclease, V
38	67.5	11.4	836	2 C97525	clpA protein (A722
39	67.5	11.4	836	2 AD2744	ATP-dependent Clp
40	67.5	11.4	963	2 T04002	hypothetical prote
41	67	11.3	140	2 T40368	hypothetical prote
42	67	11.3	247	2 AB3638	probable lipoprote
43	67	11.3	438	2 B84363	peptide chain rele
44	67	11.3	677	1 S09078	chromogranin B pre
45	67	11.3	898	2 T25167	hypothetical prote

#### ALIGNMENTS

RESULT 1  
A31768  
T-cell receptor zeta chain precursor - human  
C:Species: Homo sapiens (man)  
C>Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 09-Jul-2004  
C/Accession: A31768  
R:Weisman, A.M.; Hou, D.; Orloff, D.G.; Modi, W.S.; Senauer, H.; O'Brien, S.J.; Klausner, J.D.; Natl. Acad. Sci. U.S.A. 85, 9709-9715, 1988  
A>Title: Molecular cloning and chromosomal localization of the human T-cell receptor zeta  
A:Reference number: A31768; MUID:89071765; PMID:2974162  
A/Accession: A31768  
A:Molecule type: mRNA  
A:Residues: 1-163 <MBI>  
A/Cross-references: UNIPROT:P20963; GB:J04132; NID:G623041; PIDN:AAA60394.1; PID:G623042  
C/Keywords: phosphoprotein; T-cell receptor; transmembrane protein  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-163/Product: T-cell receptor zeta chain #status predicted <MAT>

Query Match 98.6%; Score 583; DB 2; Length 163;  
Best Local Similarity 99.1%; Pred. No. 1.9e-49;  
Matches 109; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	2	RFSRSABPPAYQGQGNOLYNELNGRREYDVLDKRGDPENGGKPRRKNPQGLYNEL	61
Db	54	KFSRSABPPAYQGQGNOLYNELNGRREYDVLDKRGDPENGGKPRRKNPQGLYNEL	113

Qy 62 QKDMAEAYSEIGMKGRGKHGDLGGLSTATDTYDALHMQALPPR 111  
Db 114 QKDMAEAYSEIGMKGRGKHGDLGGLSTATDTYDALHMQALPPR 163

RESULT 2  
A40104  
T-cell receptor CD3 zeta chain - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 20-Mar-1992 #sequence\_revision 20-Mar-1992 #text\_change 09-Jul-2004  
C/Accession: A40104; I55293  
R:Weisman, A.M.; Baniyash, M.; Hou, D.; Samelson, L.E.; Burgess, W.H.; Klausner, R.D.  
Science 239, 1018-1021, 1988  
A>Title: Molecular cloning of the zeta chain of the T cell antigen receptor.  
A:Reference number: A40104; MUID:88145643; PMID:3278377  
A/Accession: A40104  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-164 <MBI>  
A/Cross-references: UNIPROT:P24161; GB:M19729; NID:G201131; PIDN:AAA40171.1; PID:G201132  
R:Baniyash, M.; Hsu, V.M.; Seidlin, M.F.; Klausner, R.D.  
J. Biol. Chem. 264, 13252-13257, 1989  
A>Title: The isolation and characterization of the murine T cell antigen receptor zeta c  
A:Reference number: I55293; MUID:89327299; PMID:2787796  
A/Accession: I55293  
A>Status: preliminary; translated from GB/EMBL/DBJ



[illegible]

## RESULT 7

T-cell receptor CD3 eta chain precursor - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 23-Oct-1990 #sequence revision 23-Oct-1990 #text\_change 09-Jul-2004  
C/Accession: A35900; A60374; G46522  
R/Jin, Y.J.; Claydon, L.K.; Howard, F.D.; Koyasu, S.; Sieh, M.; Steinbrich, R.; Tarr, G.  
Proc. Natl. Acad. Sci. U.S.A. 87, 3319-3333, 1990  
A/Title: Molecular cloning of the CD3eta subunit identifies a CD3zeta-related product in  
A/Reference number: A35900; MUID:90239005; PMID:2139725  
A/Accession: A35900  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-206 <JIN>  
A/Cross-references: UNIPROT:P29020; GB:M33158; NID:G192488; PIDD:AA37398.1; PID:G309155  
R/Ohno, H.; Saito, T.  
Int. Immunol. 2, 1117-1119, 1990  
A/Title: CD3zeta and eta chains are produced by alternative splicing from a common gene.  
A/Reference number: A60374; MUID:91190781; PMID:2150596  
A/Accession: A60374  
A/Status: not compared with conceptual translation  
A/Molecule type: DNA  
A/Residues: 144-206 <OHNO>  
A/Jensen, J.P.; Gencarelli, C.; Hou, D.; Rellahan, B.L.; Dean, M.; Weissman, A.M.  
J. Immunol. 150, 1221-130, 1993  
A/Title: T cell antigen receptor-eta subunit. Low levels of expression and limited cross  
A/Reference number: A46522; MUID:93107707; PMID:8417118  
A/Contents: annotation  
C/Comment: The functional significance of this alternative spliced product of the CD3  
eta acids, differs widely among various mammalian species in sequence, length, and even  
A/Keywords: alternative splicing; T-cell receptor; transmembrane protein

	Query Match	64.0%	Score 378.5 ; DB 2; Length 206;
	Best Local Similarity	82.2% ; Pred. No. 1.7e-29;	
	Matches	74; Conservative 5; Mismatches 10; Indels 1; Gaps 1;	
QY	2 RFSRSAAEPAAQQGGNOLYNELNGRREBYVLDRGRDPMGSK-PRRKNPQEGLYNE	60	
	:   :	:	::   :
DQ	54 KFSRSAAETLAALADPNOLYNELNGRREYDVLEKGRARDPMGSKQGQRARNPOBGVNA	113	
	:	:	:
QY	61 LQKDMAEAYSEIGMKGERRRRGKHIDGLYQ	90	
DQ	114 LQKDMAEAYSEIGTKERRRGKHIDGLYQ	143	

## RESULT 8

regulator protein white collar 1 - *Neurospora crassa*  
C/Species: *Neurospora crassa*  
C/Date: 21-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 15-Mar-2004  
C/Accession: S69206  
R/Ballato, P.; Vittorioso, P.; Magrelli, A.; Talora, C.; Cabibbo, A.; Macino, G.  
EMBO J. 15, 1650-1657, 1996  
A/Title: White collar-1, a central regulator of blue light responses in *Neurospora*, is e  
A/Reference number: S69206, MUID:96203083, PMID:8612589  
A/Accession: S69206  
A/Status: Preliminary  
A/Molecule type: DNA  
A/Residues: 1-1154 <BAL>  
A/Cross-references: EMBL:X94300, NID:g129576, PID:g148015  
C/Genetics:  
A/Introns: 967/3  
C/Keywords: zinc finger  
;932-931/Domain: GATA-type zinc finger homology <GF>  
;932-931/Domain: GATA-type zinc finger homology <GF>

Query Match	14.6%	Score 86	DB 2	Length 1154
Best Local Similarity	24.8%	Pred. NC	2.8	
Matches 33	Conservative 19	Mismatches 49	Indels 32	Gaps 6

```

QY 9 PPAVQGGNGOLY-----NELNIGRREYVLDKRRGRDPEMGKPR-----KN 52
Db 71 PPTTNGNSTTHADVTMTSGGSDSLDIQNLDMHRRRSVPQYGGQTRRLSMFYAN 130
QY 53 PQEGLYNELQDKKAAEAYSE----IGMKGERRRKG-----HDGLYGLS--TATK 97
Db 131 PNDG-FSDYQLDNMSGNYGDMTGGMGSGHSSPYAGQINIMSDHSGYSIMSPVNMGM 189
QY 98 DTYDALHMQALPP 110
Db 190 MTEPLNINMTHSP 202

```

## RESULT 5

121095  
 Hypothetical protein F18H3.3a - *Caenorhabditis elegans*  
 C|Species: *Caenorhabditis elegans*  
 C|Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C|Accession: T21095  
 R|Coles, L.  
 submitted to the EMBL Data Library, July 1995  
 A|Reference number: 219372  
 A|Accession: T21095  
 A|Status: preliminary; translated from GB/EMBL/DBJ  
 A|Molecule type: DNA  
 A|Residues: 1-662 <MIL>  
 A|Cross-references: UNIPROT:Q19579, EMBL:Z50110, PIDB:CAA90444.1, GSPDB:GN00028, CESP:FL18H3  
 A|Experimental source: clone F18H3  
 C|Genetics:  
 A|Gene: CESP:F18H3.3a  
 A|Map position: X  
 A|Introns: 111/1, 215/2, 469/3, 552/3  
 C|Superfamily: polyadenylation-binding protein; ribonucleoprotein repeat homology

[illegible]

**RESULT 10**

probable ATP-dependent RNA helicase Ddbp45A - fruit fly (*Drosophila melanogaster*)  
N:Alternate names: DEAD-box protein  
C:Species: *Drosophila melanogaster*  
C:Date: 31-Dec-1993 #sequence\_revision 02-Aug-1994 #text\_change 09-Jul-2004  
C:Accession: S38329; S39749  
R:Lavoie, C.A.; Harvey, M.; Lasko, P.F.  
Biochim. Biophys. Acta 1216, 140-144, 1993  
A>Title: Ddbp45A encodes a *Drosophila* DEAD box protein with similarity to a putative yeast  
A:Reference number: S38329; MUID:94032476; PMID:7692973  
A:Accession: S38329  
A:Molecule type: DNA  
A:Residues: 1-527 <LAV1>  
A:Cross-references: UNIPROT:007886; GB:U13612  
A:Notes: The authors translated the codon TGC for residue 211 as Ser, GAG for residue 212  
as Gln, GGT for residue 225 as Arg, GTC for residue 228 as Leu, TCG for residue 229 as

Riavoie, C.A.; Harvey, M.; Iasko, P.F.  
 submitted to the EMBL Data Library, June 1993  
 A:Reference number: S39749  
 A:Accession: S39749  
 A:Molecule type: DNA  
 A:Residues: 1-328, 'I', 328-406, 'C', 408-492, 'Y', 494-527 <LAVZ>  
 A:Cross-references: EMBL:LI3612; NID:g304670; PIDD:AAAL6339.1; PID:g304671  
 C:Genetics:  
 A:Gene: DBP45A  
 A:Cross-references: FlyBase:FBgn0010220  
 A:Introns: 26/1; 206/3; 262/2; 382/1; 433/2  
 C:Keywords: ATP, nucleotide binding; P-loop  
 F:51-58/Region: nucleotide-binding motif A (P-loop)  
 F:153-158/Region: nucleotide-binding motif B  
 F:157-160/Region: DEAD motif

Query Match 12.5%; Score 76.5; DB 2; Length 527;  
 Best Local Similarity 23.5%; Pred. No. 9.7;  
 Matches 32; Conservative 19; Mismatches 28; Indels 57; Gaps 6;

QY 4 SRSAPPAYQGGONQLYNELNGRRREYDVLDKR-----RGRDPMGSK 47  
 DB 373 TKLTREHIDPRMVERIFMQVNVTRSEMQLDNNPDERAQNRYRKTMTMEGDPDQ--- 429  
 QY 48 PRKRNPQEGLYNELQKDKMAE-----AYSEIG--MKGR-----RRG 82  
 DB 430 -----MEALYRKQKQKRLKRIKRLQHAESAASEKALLQDERFKSVDSARPEKKG 483  
 QY 83 KGHDLGYOGISTATKCD 98  
 DB 484 KGR-----SRATQEE 492

RESULT 11  
 T14305  
 probable late embryogenesis abundant protein - carrot (fragment)  
 N:Alternate names: LEA protein  
 C:Species: Daucus carota (carrot)  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
 C:Accession: T14305  
 Riabin, X.; Hwang, G.J.; Zimmerman, J.L.  
 submitted to the EMBL Data Library, January 1996  
 A:Description: Isolation and characterization of a diverse set of genes from carrot soma  
 A:Reference number: Z17968  
 A:Accession: T14305  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-136 <LIN>  
 A:Cross-references: UNIPROT:Q39681; EMBL:U47096; NID:gl276968; PID:gl276969  
 A:Experimental source: strain Danver Half-Long

Query Match 12.5%; Score 74; DB 2; Length 136;  
 Best Local Similarity 26.8%; Pred. No. 3.7;  
 Matches 22; Conservative 14; Mismatches 38; Indels 8; Gaps 2;

QY 29 EERYDVLDKRGRDPMGSKPRRKNPQEGLYNELQ-----KDKMAEAYSEIGMKERRRG 82  
 DB 1 EADYDTGKADKQKVDLTGKAKDAEAKA--NEASDITGKADKQKAYEAAEKGNBAEWMTG 58  
 QY 83 KGHDLGYOGISTATKCDYDALH 104  
 DB 59 KADKAYETASAKKCANEAAY 80

RESULT 12  
 T13828  
 CREB-binding protein homolog - fruit fly (Drosophila melanogaster)  
 C:Species: Drosophila melanogaster  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
 C:Accession: T13828  
 R;Altman, H.; Chen, Y.; Dai, P.; Hou, D.X.; Nonaka, M.; Smolik, S.M.; Armstrong, S.; Gd  
 Nature 386, 735-738, 1997  
 A:Title: Drosophila CBP is a co-activator of cubitus interruptus in hedgehog signalling.

A:Reference number: Z17785; MUID:97263578; PMID:9109493  
 A:Accession: T13828  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-3190 <AKI>  
 A:Cross-references: UNIPROT:O01368; EMBL:U88570; NID:g1916929; PID:g1916930; PIDD:AA8530;  
 A:Genetics:  
 A:Gene: gsaB  
 A:Map position: X  
 F:1723-1780/Domain: bromodomain homology <BRO>

Query Match 12.5%; Score 74; DB 2; Length 3190;  
 Best Local Similarity 26.2%; Pred. No. 1.3e+02;  
 Matches 27; Conservative 23; Mismatches 35; Indels 18; Gaps 5;

QY 9 PPAYQGGONQLYNELNGRRREYDVLDKRGRDPMGSKPRRKNPQEGLYNELQKDKMAE 68  
 DB 2881 PQNPQQ 2930  
 QY 69 AYSEIGMKGRERRRGKGDGLYOGISTATKCDYDALH-QALPP 110  
 DB 2931 L-----MQQQQGGGPGG--SGLPTRPMPNPLGWLQSLPP 2966

RESULT 13  
 S26183  
 general stress protein gsaB - Bacillus subtilis  
 N:Alternate names: glucose starvation inducible B  
 C:Species: Bacillus subtilis  
 C:Date: 22-Nov-1993 #sequence\_revision 02-Jun-1994 #text\_change 09-Jul-2004  
 C:Accession: S26183; H69637; S21381  
 R;Mueller, J.P.; Bukusoglu, G.; Sonenshein, A.L.  
 J. Bacteriol. 174, 4361-4373, 1992  
 A:Title: Transcriptional regulation of Bacillus subtilis glucose starvation-inducible gsaB  
 A:Reference number: S26183; MUID:92325021; PMID:1378051  
 A:Accession: S26183  
 A:Molecule type: DNA  
 A:Residues: 1-123 <MUB>  
 A:Cross-references: UNIPROT:P26907; EMBL:X56680; NID:g39933; PIDD:CAA4009.1; PID:g39934  
 A:Note: the authors translated the codon AAA for residue 6 as leu  
 R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azavedo, V.; Beretx  
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi  
 A.; Ehrlich, S.D.; Emmergon, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, B.  
 Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Funai, S.; Galizzi, A.; Gallert  
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtschappel, S.; Hosono, S.; Hullo, M.F.  
 Koeter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauele  
 Y., M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
 Rieger, M.; Rivolta, S.; Schreier, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot  
 A:Authors: Schlögl, S.; Schreier, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot  
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunori, A.; Yoshida, K  
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A:Reference number: A69580; MUID:98044033; PMID:9384377  
 A:Accession: H69637  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-123 <KUN>  
 A:Cross-references: GB:299106; GB:AL009126; NID:g2632653; PIDD:CA812247.1; PID:g2632740  
 A:Experimental source: strain 168  
 C:Comment: This protein of unknown function is induced by glucose starvation or phosphate

Query Match 12.4%; Score 73.5; DB 1; Length 123;  
 Best Local Similarity 31.4%; Pred. No. 3.7;  
 Matches 22; Conservative 7; Mismatches 28; Indels 13; Gaps 3;

Oy	45	G G K R R N N P O B G L V N E L - - - - -	- O K D K M A E A Y S E I G M K G E R R R K G H D - G L Y O O L - - - -	92
		: : : :		
Db	16	G G E T T S N K M H K F Y Q E I G C K G E A T S N K M H K F Y Q E I G E K G E A T S N K M D K F Y Q E I G E K		75
Oy	93	- S T A T K D T Y P	101	
		: : : :		
Db	76	G G E A T S E N M D	85	

Db 140 LVPHHDLVASSVGSYSALIKENEDTV 168

Search completed: March 7, 2005, 07:21:03

Job time : 11.1499 secs

**RESULT 14**

rare lipoprotein A precursor [imported] - Yersinia pestis (strain CO92)  
 C:Species: Yersinia pestis  
 C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
 C:Accession: AF0317  
 R:Parkhill, J., Wren, B.W., Thomson, N.R., Titball, R.W., Holden, M.T.G., Prentice, M.B.  
 demo-Tarraga, A.M., Chillingworth, T., Cronin, A., Davies, R.M., Davis, P., Dougan, G.,  
 Hill, M., Rutherford, K., Simmonds, M., Skelton, J., Stevens, K., Whitehead, S., Barrrell,  
 Nature 413, 523-527, 2001  
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
 A:Reference number: AB0001; PMID:11586360  
 A:Accession: AF0317  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-360 <KUR>  
 A:Cross-references: UNIPROT:Q8ZDG6; GB:AL590842; PIDN:G15980589; GSPDB:  
 C:Genetics:  
 A:Gene: rlpA  
 A:Superfamily: rlpA lipoprotein

Query Match	12.4%	Score	73.5	DB	2	Length	360
Best Local Similarity	25.9%	Pred. No.	12				
Matches	29	Conservative	17	Mismatches	33	Indels	33
						Gaps	6

```
QY      8 EPPAYOQGQNLYN----ELNLGRREBYDLKCRGRDPMGKKPR--KNPOEGLYNEL    61
        :||| ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db     21 QRPACQQQVVOQTYSGPVEIG-GAPRPEREPFNPNVNQDYKVGSGYRIIKDPQN-----    73
```

```

Oy      62 QKDRAEAYSEIGMK--GERRRGKHGDLVQGLSTATKDTYDALHMOALPP 110
          : : : : : : : : : : : : : : : : : : : : : : : :
Db      74 -----FSQIGLASSYGEARGN-----TTATCEIFDPNALTAHP 108

```

## RESULT 15

**H70415**

conserved hypothetical protein aq\_1336 - Aquifex aeolicus  
C/Species: Aquifex aeolicus

C:\Date: 08-May-1998 #sequence\_ revision 08-May-1998 #text\_ change 09-Jul-2004  
C:\Access\Bion: H70415

R, Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.

Nature 392, 353-358, 1998

**A/Title:** The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.

A;Reference number: A70300; MUID:98196666; PMID:9537320

A;Accession: H/0415

A/status: preliminary; nucleic acid sequence not shown; translation not shown

A; MOLECULE TYPE: DNA

A;KCBIAUEB: 1-3/4 &lt;AUE&gt;

A: Experimental source: strain VFE  
A/C1088-1  
EIDENCEB: UNIFR01:08/333; GB:AE00V/33; MID:g2383/49; FIDN:MACU/322.1; FID:g2383/49

C:\Genetics:

A:Gene: ag 1336

1

Query Match	12.4%	Score 73.5;	DB 2;	Length 374;
Best Local Similarly	29.2%	Pred. No. 13;		
Matches 26; Conservative	21;	Mismatches 37;	Indels 5;	Gaps 4;

19 LYNELNIGR-REYDVLDRGRGDPBMGCKPRKNPQEGLYNELCK-DKMAEAYSSEIGM 75

```

Db      81  LYNERFLNKLKERERKILKGYRPSHAGQAYE
          |||  :  :  :  :  :  :  :  :
Qy      76  KGERRGKGHDGLYOGISTATKDT-YDAL 103
          :  ||  :  :  :  :  :  :

```

**This Page Blank (uspto)**